Appl. Comput. Harmon. Anal. 39 (2015) 1–20

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Greedy signal space methods for incoherence and beyond

Raja Giryes [a,*], Deanna Needell [b]

[a] Department of Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel
[b] Department of Mathematical Sciences, Claremont McKenna College, Claremont, CA 91711, USA

A R T I C L E I N F O

Article history:
Received 12 September 2013
Received in revised form 31 March 2014
Accepted 21 July 2014
Available online 27 July 2014
Communicated by Henrique Malvar

MSC:
94A20
94A12
62H12

Keywords:
Algorithms
Approximation
Compressed Sensing
Restricted Isometry Property
Sparse modeling
CoSaMP
Sparse approximation
Signal space methods
Coherent dictionaries

A B S T R A C T

Compressive sampling (CoSa) has provided many methods for signal recovery of signals compressible with respect to an orthonormal basis. However, modern applications have sparked the emergence of approaches for signals not sparse in an orthonormal basis but in some arbitrary, perhaps highly overcomplete, dictionary. Recently, several "signal-space" greedy methods have been proposed to address signal recovery in this setting. However, such methods inherently rely on the existence of fast and accurate projections which allow one to identify the most relevant atoms in a dictionary for any given signal, up to a very strict accuracy. When the dictionary is highly overcomplete, no such projections are currently known; the requirements on such projections do not even hold for incoherent or well-behaved dictionaries. In this work, we provide an alternate analysis for signal space greedy methods which enforce assumptions on these projections which hold in several settings including those when the dictionary is incoherent or structurally coherent. These results align more closely with traditional results in the standard CoSa literature and improve upon previous work in the signal space setting.

1. Introduction

In many signal and image processing applications we encounter the following problem: recovering an original signal $\mathbf{x} \in \mathbb{R}^d$ from a set of noisy measurements

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{M} \in \mathbb{R}^{m \times d}$ is a known linear operator and $\mathbf{e} \in \mathbb{R}^d$ is additive bounded noise, i.e. $\|\mathbf{e}\|_2^2 \le \varepsilon^2$. In many cases such as those in Compressive Sampling (CoSa) [1], we have $m \ll d$ and thus (1) has infinitely

* Corresponding author.
E-mail addresses: raja@cs.technion.ac.il (R. Giryes), dneedell@cmc.edu (D. Needell).

http://dx.doi.org/10.1016/j.acha.2014.07.004
1063-5203/

many solutions. To make the problem well-posed we rely on additional priors for the signal $\mathbf{x}$, such as *sparsity*.

The sparsity assumption provides two main models, termed the *synthesis* and *analysis* models [2]. The synthesis model, which has received great attention in the past decade, assumes that $\mathbf{x}$ has a k-sparse representation $\boldsymbol{\alpha}$ under a given dictionary $\mathbf{D} \in \mathbb{R}^{d\times n}$ [3]. In other words, there exists a vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ such that $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$ and $\|\boldsymbol{\alpha}\|_0 \le k$, where $\|\boldsymbol{\alpha}\|_0 = |\text{supp}(\boldsymbol{\alpha})|$ denotes the ℓ_0 pseudo-norm. Under the synthesis model assumption we can recover $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$ by solving

$$\operatorname*{argmin}_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|_2 \le \varepsilon. \tag{2}$$

Since solving (2) is an NP-complete problem in general [4], approximation techniques are required for recovering $\mathbf{x}$. One strategy uses relaxation, replacing the ℓ_0 with the ℓ_1 norm, resulting in the ℓ_1-synthesis problem

$$\hat{\boldsymbol{\alpha}}_{\ell_1} = \operatorname*{argmin}_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|_2 \le \varepsilon. \tag{3}$$

The study of these types of synthesis programs has largely relied on properties like the *Restricted Isometry Property* (RIP) [5], which states that

$$(1-\delta_k)\|\mathbf{x}\|^2 \le \|\mathbf{M}\mathbf{x}\|^2 \le (1+\delta_k)\|\mathbf{x}\|^2 \quad \text{for all } k\text{-sparse } \mathbf{x},$$

for some small enough constant $\delta_k < 1$.

If the matrix $\mathbf{D}$ is unitary and the vector $\mathbf{x}$ has a k-sparse representation $\boldsymbol{\alpha}$, then when $\mathbf{M}$ satisfies the RIP with $\delta_{2k} < \delta_{\ell_1}$, the program (3) accurately recovers the signal,

$$\|\hat{\mathbf{x}}_{\ell_1} - \mathbf{x}\|_2 \le C_{\ell_1}\varepsilon, \tag{4}$$

where $\hat{\mathbf{x}}_{\ell_1} = \mathbf{D}\hat{\boldsymbol{\alpha}}_{\ell_1}$, C_{ℓ_1} is a constant greater than $\sqrt{2}$ and δ_{ℓ_1} ($\simeq 0.4652$) is a constant [6–8]. This result also implies perfect recovery in the absence of noise. It was extended also for incoherent redundant dictionaries [9].

An alternative approach to approximating (2) is to use a greedy strategy. Recently introduced methods that use this strategy are the CoSaMP [10], IHT [11], and HTP [12] methods. Greedy methods iteratively identify elements of the support of the signal, and once identified, use a simple least-squares to recover the signal. These methods were shown to have guarantees in the form of (4) under the assumption of the RIP. However, such results hold only when $\mathbf{D}$ is orthonormal, and do not hold for general dictionaries $\mathbf{D}$. Recently, the greedy approaches have been adapted to this setting. For example, the *Signal Space CoSaMP* method [13] adapts CoSaMP to the setting of arbitrary dictionaries. A slight modification[1] of this method is shown in Algorithm 1. In the algorithm, the subscript T denotes the restriction to elements (columns) indexed in T. The function $\mathscr{S}_k(\mathbf{y})$ returns the support of the best k-sparse representation of $\mathbf{y}$ in the dictionary $\mathbf{D}$, and $\mathbf{P}_T$ denotes the projection onto that support.

In [13], the authors analyze this CoSaMP variant under the assumption of the $\mathbf{D}$-*RIP* [14], which states[2]

$$(1-\delta_k)\|\mathbf{D}\boldsymbol{\alpha}\|^2 \le \|\mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|^2 \le (1+\delta_k)\|\mathbf{D}\boldsymbol{\alpha}\|^2 \quad \text{for all } k\text{-sparse } \boldsymbol{\alpha}. \tag{5}$$

They prove that under this assumption, if one has access to projections $\mathscr{S}_k$ which satisfy

[1] Here we use two separate support selection schemes, whereas the original Signal Space CoSaMP method uses one.

[2] By abuse of notation we denote both the RIP and the $\mathbf{D}$-RIP constants by δ_k. It will be clear from the context to which one we refer at each point in the article.